



Detecting abrupt changes of the long-range dependence or the self-similarity of a Gaussian process

Jean-Marc Bardet, Imen Kammoun

► To cite this version:

Jean-Marc Bardet, Imen Kammoun. Detecting abrupt changes of the long-range dependence or the self-similarity of a Gaussian process. 2007. hal-00195089v2

HAL Id: hal-00195089

<https://hal.science/hal-00195089v2>

Preprint submitted on 26 Apr 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Detecting abrupt changes of the long-range dependence or the self-similarity of a Gaussian process

Jean-Marc Bardet^a, Imen Kammoun^a

^a *Université Paris 1, SAMOS-MATISSE-CES, 90 rue de Tolbiac, 75013 Paris, France*

Received *****, accepted after revision +++++

Presented by $\mathbb{L}\mathbb{L}\mathbb{L}\mathbb{L}\mathbb{L}$

Abstract

In this paper, an estimator of m instants (m is known) of abrupt changes of the parameter of long-range dependence or self-similarity is proved to satisfy a limit theorem with an explicit convergence rate for a sample of a Gaussian process. In each estimated zone where the parameter is supposed not to change, a central limit theorem is established for the parameter's (of long-range dependence, self-similarity) estimator and a goodness-of-fit test is also built. *To cite this article: J.M. Bardet, I. Kammoun, C. R. Acad. Sci. Paris, Ser. I 340 (2007).*

Résumé

Détection de ruptures du paramètre de longue mémoire, d'autosimilarité pour des processus gaussiens. Dans ce papier, pour une trajectoire d'un processus gaussien, un estimateur des m points de ruptures (m est supposé connu) du paramètre de longue mémoire ou d'autosimilarité est construit et on montre qu'il vérifie un théorème limite avec une vitesse de convergence explicite. Dans chaque zone (estimée) où ce paramètre est constant, un estimateur de ce paramètre vérifie un théorème limite centrale et un test d'ajustement est galemment mis en place. *Pour citer cet article : J.M. Bardet, I. Kammoun, C. R. Acad. Sci. Paris, Ser. I 340 (2007).*

Version française abrégée Le problème de détection de points de ruptures moyennant la minimisation d'une fonction de contraste donnée a été étudié depuis le milieu des années 1990 dans le cadre de processus à longue mémoire (voir par exemple [9], [11], [12], [13] et [14]). De ces approches, certaines ont été associées à un cadre paramétrique, telle que la détection de rupture selon la moyenne et/ou la variance, d'autres traitées dans un cadre non paramétrique (comme la détection de ruptures selon la distribution ou le spectre). Dans la littérature, différents auteurs ont également proposé des statistiques de test de l'hypothèse que le paramètre est inchangé contre le fait que le paramètre de longue mémoire varie en fonction du temps (voir par exemple [2], [7], [10]). À notre connaissance, le cadre semi-paramétrique de détection de changements en longue mémoire ou en autosimilarité n'a t trait que dans [13] partir d'une

Email addresses: jean-marc.bardet@univ-paris1.fr (Jean-Marc Bardet), imen.kammoun@univ-paris1.fr (Imen Kammoun).

technique base sur le priodogramme.

Notre approche est fonde sur l'analyse par ondelettes, ce qui prsente plusieurs avantages : c'est une technique non-paramtrique applicable pour des processus trs gnraux, robuste aux tendances polynomiales et, au moins dans le cadre gaussien, s'accompagnant de tests d'adquation de type χ^2 simples et intressants utiliser. Ainsi, un estimateur des m points de ruptures ($m \in \mathbb{N}^*$, suppos connu) de la longue mémoire ou d'auto-similarité est conçu pour un échantillon de processus gaussien en se basant sur l'analyse par ondelettes, ce qui permet ensuite de mettre en place des tests d'adquation. Pour ce type de processus, cette méthode a été proposée pour la première fois dans [8], puis développée par exemple dans [1]. La convergence des estimateurs basés sur les ondelettes a été étudiée dans le cas du mouvement brownien fractionnaire (FBM) dans [3], et dans un cadre semi-paramétrique général de processus gaussiens stationnaires à longue mémoire par [15] et [5].

Ici le principe de l'estimation du paramètre de longue mémoire ou d'auto-similarité est le suivant : dans chaque zone où il n'y a pas de changement, ce paramètre peut être estimé à partir d'une log-log régression de la variance des coefficients d'ondelettes sur plusieurs échelles choisies (voir (1)). Une fonction de contraste définie par la somme des carrés des distances entre ces points et les droites d'ajustement, dans les $m + 1$ zones possibles détectées, est minimisée (voir (2)), donnant un estimateur des points de ruptures (voir (3)). Sous certaines hypothèses générales, on montre qu'il vérifie un théorème limite avec une vitesse de convergence explicite (voir Theorem 1.1). Dans chacune des zones détectées, les paramètres de longue mémoire, (ou d'auto-similarité) peuvent être estimés, tout d'abord avec la regression des moindres carrés ordinaires (OLS), puis par une regression des moindres carrés pseudo-généralisés (FGLS). Un théorème de la limite centrale est établi pour chacun des deux estimateurs (voir Theorem 1.2 et Proposition 1.1 ci-dessous) et des intervalles de confiance peuvent être calculés. L'estimateur FGLS offre deux avantages : d'une part, sa variance asymptotique est plus petite que celle de l'estimateur OLS, et d'autre part, il permet la construction d'un test d'ajustement très simple basé sur le carré des distances entre les points (d'abscisse, le logarithme d'une échelle choisie et d'ordonnée, le logarithme de la variance empirique des coefficients d'ondelettes pour cette échelle) et les droites de régression pseudo-généralisée correspondantes (voir (8)). La convergence vers une distribution du Chi-deux de ce test est établie dans le Theorem 1.3. Deux cas particuliers de processus gaussiens sont ensuite étudiés dans la section 2. En premier lieu, on s'intéresse aux séries chronologiques stationnaires longue mémoire avec un paramètre de Hurst constant par morceaux. On se place dans un cadre semi-paramétrique contenant par exemple les FGN et les processus FARIMA (voir Figure 1). En second lieu, le cas d'un processus à accroissements stationnaires et autosimilaire par morceaux est traité, ce qui revient à considérer des successions de FBM ayant des exposants de Hurst distincts (voir Figure 1). Pour ces deux exemples, les vitesses de convergence explicites des différents estimateurs et tests sont données et des simulations montrent leurs qualités (voir Table 1). D'autres simulations, preuves des thormes ainsi qu'un exemple plus gnral de dttection de ruptures dans le cadre de processus gaussien localement fractionnaire sont dtaills dans [6].

1. Assumptions and main results

Let $(X_t)_{t \in \mathbb{N}}$ be a Gaussian process and assume that (X_0, X_1, \dots, X_N) is known. In the sequel, X will be a piecewise stationary long memory time series or a piecewise self-similar time series having stationary increments. Consider $\psi : \mathbb{R} \rightarrow \mathbb{R}$ a function called "the mother wavelet". For $(a, b) \in \mathbb{R}_+^* \times \mathbb{R}$, the wavelet coefficient of X for the scale a and the shift b is $d_X(a, b) := \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \psi(\frac{t-b}{a}) X(t) dt$. When only a discretized path of X is available, approximations $e_X(a, b)$ are only computable :

$$e_X(a, b) := \frac{1}{\sqrt{a}} \sum_{p=1}^N \psi\left(\frac{p-b}{a_N}\right) X_p \quad \text{for } (a, b) \in \mathbb{R}_+^* \times \mathbb{N}.$$

Assume that there exist $m \in \mathbb{N}$ (the number of abrupt changes) and

- $0 = \tau_0^* < \tau_1^* < \dots < \tau_m^* < \tau_{m+1}^* = 1$ (unknown parameters);
- two families $(\alpha_j^*)_{0 \leq j \leq m} \in \mathbb{R}^{m+1}$ and $(\beta_j^*)_{0 \leq j \leq m} \in (0, \infty)^{m+1}$ (unknown parameters);
- a sequence of "scales" $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ (chosen) satisfying $a_n \geq a_{\min}$, with $a_{\min} > 0$,

such that for $j = 0, 1, \dots, m$ and $k \in D_N^*(j) \subset [[N\tau_j^*], [N\tau_{j+1}^*]]$,

$$\mathbb{E}[e_X^2(a_N, k)] \sim \beta_j^*(a_N)^{\alpha_j^*} \text{ when } N \rightarrow \infty.$$

A piecewise sample variance can be the appropriated estimator of such power law. Thus, define

$$S_k^{k'}(a_N) := \frac{a_N}{k' - k} \sum_{p=[k/a_N]}^{[k'/a_N]-1} e_X^2(a_N, a_N p) \quad \text{for } 0 \leq k < k' \leq N.$$

Now set $0 < r_1 < \dots < r_\ell$ with $\ell \in \mathbb{N}^*$, and assume that a multidimensional central limit theorem can be established for $\left(\log(S_k^{k'}(r_i a_N)) \right)_{1 \leq i \leq \ell}$, when $[N\tau_j^*] \leq k < k' \leq [N\tau_{j+1}^*]$, i.e.

$$\sqrt{\frac{k' - k}{a_N}} \left(\log(S_k^{k'}(r_i a_N)) - \log(\beta_j^*) - \alpha_j^* \log(r_i a_N) \right)_{1 \leq i \leq \ell} \xrightarrow[N \rightarrow \infty]{|k' - k| \rightarrow \infty} \mathcal{N}(0, \Gamma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell)), \quad (1)$$

with $\Gamma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell) = (\gamma_{pq}^{(j)})_{1 \leq p, q \leq \ell}$ a $(\ell \times \ell)$ matrix not depending on N such that $\alpha \mapsto \Gamma^{(j)}(\alpha, r_1, \dots, r_\ell)$ is a continuous function and a positive matrix for all α . Define a contrast function

$$U_N((\alpha_j)_{0 \leq j \leq m}, (\beta_j)_{0 \leq j \leq m}, (k_j)_{1 \leq j \leq m}) = \sum_{j=0}^m \sum_{i=1}^{\ell} \left(\log(S_{k_j}^{k_{j+1}}(r_i a_N)) - (\alpha_j \log(r_i a_N) + \log \beta_j) \right)^2 \quad (2)$$

with $(\alpha_j)_{0 \leq j \leq m} \in A^{m+1} \subset \mathbb{R}^{m+1}$, $(\beta_j)_{0 \leq j \leq m} \in B^{m+1} \subset (0, \infty)^{m+1}$, $0 = k_0 < k_1 < \dots < k_m < k_{m+1} = N$, $(k_j)_{1 \leq j \leq m} \in K_m(N) \subset \mathbb{N}^m$. The vector of estimated parameters $\hat{\alpha}_j$, $\hat{\beta}_j$ and \hat{k}_j (and therefore $\hat{\tau}_j$) is the vector which minimizes this contrast function in $A^{m+1} \times B^{m+1} \times K_m(N)$, i.e.,

$$((\hat{\alpha}_j)_{0 \leq j \leq m}, (\hat{\beta}_j)_{0 \leq j \leq m}, (\hat{k}_j)_{1 \leq j \leq m}) := \text{Argmin} \left\{ U_N((\alpha_j)_{0 \leq j \leq m}, (\beta_j)_{0 \leq j \leq m}, (k_j)_{1 \leq j \leq m}) \right\} \quad (3)$$

$$\hat{\tau}_j := \hat{k}_j / N \quad \text{for } 1 \leq j \leq m. \quad (4)$$

For a given $(k_j)_{1 \leq j \leq m}$, it is obvious that $(\hat{\alpha}_j)_{0 \leq j \leq m}$ and $(\log \hat{\beta}_j)_{0 \leq j \leq m}$ are obtained from a log-log regression of $(S_{k_j}^{k_{j+1}}(r_i a_N))_i$ onto $(r_i a_N)_i$, i.e.

$$\begin{pmatrix} \hat{\alpha}_j \\ \log \hat{\beta}_j \end{pmatrix} = (L'_1 \cdot L_1)^{-1} L'_1 \cdot Y_{k_j}^{k_{j+1}} \quad \text{with } Y_{k_j}^{k_{j+1}} := (\log(S_{k_j}^{k_{j+1}}(r_i \cdot a_N)))_{1 \leq i \leq \ell}, \quad L_{a_N} := \begin{pmatrix} \log(r_1 a_N) & 1 \\ \vdots & \vdots \\ \log(r_\ell a_N) & 1 \end{pmatrix}.$$

Therefore $(\hat{k}_j)_{1 \leq j \leq m} = \text{Argmin} \left\{ U_N((\hat{\alpha}_j)_{0 \leq j \leq m}, (\hat{\beta}_j)_{0 \leq j \leq m}, (k_j)_{1 \leq j \leq m}), (k_j)_{1 \leq j \leq m} \in K_m(N) \right\}$.

Remark 1 In this paper, m is supposed to be known. However, if m is unknown, as in [12] or [13], a penalized contrast $\tilde{U}_{m,N} = U_N + \beta_N \times m$ (with β_N an appropriated sequence converging to 0) can be used instead of U_N , and by adding a minimization in m , an estimator \hat{m} of m could be also deduced.

In this paper, parameters (α_j^*) are supposed to satisfied abrupt changes:

Assumption C : Parameters (α_j^*) are such that $|\alpha_{j+1}^* - \alpha_j^*| \neq 0$ for all $j = 0, 1, \dots, m-1$.

Theorem 1.1 Define $\underline{\tau}^* := (\tau_1^*, \dots, \tau_m^*)$, $\hat{\underline{\tau}} := (\hat{\tau}_1, \dots, \hat{\tau}_m)$ and $\|\underline{\tau}\|_m := \max(|\tau_1|, \dots, |\tau_m|)$. Let $\ell \in \mathbb{N} \setminus \{0, 1, 2\}$. If Assumption C and relation (1) holds with $(\alpha_j^*)_{0 \leq j \leq m}$ such that $\alpha_j^* \in [a, a']$ and $a < a'$ for all $j = 0, \dots, m$, then if $a_N^{1+2(a'-a)} N^{-1} \xrightarrow{N \rightarrow \infty} 0$, for all $(v_n)_n$ satisfying $v_N \cdot a_N^{1+2(a'-a)} N^{-1} \xrightarrow{N \rightarrow \infty} 0$,

$$\mathbb{P}\left(v_N \|\underline{\tau}^* - \hat{\underline{\tau}}\|_m \geq \eta\right) \xrightarrow{N \rightarrow \infty} 0 \quad \text{for all } \eta > 0. \quad (5)$$

Remark 2 The proof of this result is provided in [6]. Unfortunately, the rate of convergence of $\|\underline{\tau}^* - \hat{\underline{\tau}}\|_m$ is only $v_N = N^\alpha$ with $0 < \alpha < 1$ and not N as, for instance, in [12] and [13]. However the context is not the same: in these papers, the contrast is directly computed from N values of $(X)_i$ which do not change following $\hat{\tau}$. Here, the contrast is computed from only $(m+1)\ell$ values of S_N which change following $\hat{\tau}$. The rate of convergence N can not be reached in such a context (simulations show also this property). This is certainly a drawback of your method, which hopefully does not change the rate of convergence of parameters (α_j) and (β_j) .

For $j = 0, 1, \dots, m$, the log-log regression of $(S_{\hat{k}_j}^{\hat{k}_{j+1}}(r_i a_N))_{1 \leq i \leq \ell}$ onto $(r_i a_N)_{1 \leq i \leq \ell}$ provides estimators of α_j^* and β_j^* . However, if τ_j converges to τ_j^* , $\hat{k}_j = N \cdot \hat{\tau}_j$ does not converge to k_j^* , and therefore $\mathbb{P}([\hat{k}_j, \hat{k}_{j+1}] \subset [k_j^*, k_{j+1}^*])$ does not tend to 1. So, define \tilde{k}_j and \tilde{k}'_j such that $\tilde{k}_j = \hat{k}_j + \frac{N}{v_N}$ and $\tilde{k}'_j = \hat{k}_{j+1} - \frac{N}{v_N}$. From (5) with $\eta = 1/2$, $\mathbb{P}([\tilde{k}_j, \tilde{k}'_j] \subset [k_j^*, k_{j+1}^*]) \xrightarrow{N \rightarrow \infty} 1$. Then,

Theorem 1.2 Let $\Theta_j^* := \begin{pmatrix} \alpha_j^* \\ \log \beta_j^* \end{pmatrix}$ and $\tilde{\Theta}_j := (L'_1 L_1)^{-1} L'_1 Y_{\tilde{k}_j}^{\tilde{k}'_j} = \begin{pmatrix} \tilde{\alpha}_j \\ \log \tilde{\beta}_j \end{pmatrix}$. Under the same assumptions as in Theorem 1.1, for $j = 0, \dots, m$, with $\Sigma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell) := (L'_1 L_1)^{-1} L'_1 \Gamma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell) L_1 (L'_1 L_1)^{-1}$,

$$\sqrt{\frac{N(\tau_{j+1}^* - \tau_j^*)}{a_N}} (\tilde{\Theta}_j - \Theta_j^*) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \Sigma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell)) \quad (6)$$

A second estimator of Θ_j^* can be obtained from feasible generalized least squares (FGLS) estimation. Indeed, the asymptotic covariance matrix $\Gamma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell)$ can be estimated by the matrix $\tilde{\Gamma}^{(j)} := \Gamma^{(j)}(\tilde{\alpha}_j, r_1, \dots, r_\ell)$ and $\tilde{\Gamma}^{(j)} \xrightarrow[N \rightarrow \infty]{\mathcal{P}} \Gamma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell)$. Then, the FGLS estimator $\bar{\Theta}_j$ of Θ_j^* is defined from the minimization among all Θ of the following squared distance,

$$\|Y_{\tilde{k}_j}^{\tilde{k}'_j} - L_{a_N} \cdot \Theta\|_{\tilde{\Gamma}^{(j)}}^2 = (Y_{\tilde{k}_j}^{\tilde{k}'_j} - L_{a_N} \Theta)' \cdot (\tilde{\Gamma}^{(j)})^{-1} \cdot (Y_{\tilde{k}_j}^{\tilde{k}'_j} - L_{a_N} \Theta).$$

and therefore define $\bar{\Theta}_j := (L'_1 (\tilde{\Gamma}^{(j)})^{-1} L_1)^{-1} L'_1 (\tilde{\Gamma}^{(j)})^{-1} Y_{\tilde{k}_j}^{\tilde{k}'_j}$.

Proposition 1.1 Under the same assumptions as in Theorem 1.2, for $j = 0, \dots, m$

$$\sqrt{\frac{N(\tau_{j+1}^* - \tau_j^*)}{a_N}} (\bar{\Theta}_j - \Theta_j^*) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, M^{(j)}(\alpha_j^*, r_1, \dots, r_\ell)) \quad (7)$$

with $M^{(j)}(\alpha_j^*, r_1, \dots, r_\ell) := (L'_1 (\Gamma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell))^{-1} L_1)^{-1} \leq \Sigma^{(j)}(\alpha_j^*, r_1, \dots, r_\ell)$ (for the order's relation between positive symmetric matrix).

Therefore $\bar{\alpha}_j$ is more accurate than $\tilde{\alpha}_j$ for estimating α_j^* when N is large enough. For $j = 0, \dots, m$, let $T^{(j)}$ be the FGLS distance between points $(\log(r_i a_N), \log(S_{\tilde{k}_j}^{\tilde{k}'_j}))_{1 \leq i \leq \ell}$ and the FGLS regression line. The following theorem describes the asymptotic behavior of a goodness-of-fit test on each segment $[\tilde{k}_j, \tilde{k}'_j]$:

Theorem 1.3 Under the same assumptions as in Theorem 1.1, for $j = 0, \dots, m$

$$T^{(j)} = \frac{N(\tau_{j+1}^* - \tau_j^*)}{a_N} \|Y_{\tilde{k}_j}^{\tilde{k}'} - L_{a_N} \bar{\Theta}_j\|_{\tilde{\Gamma}^{(j)}}^2 \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \chi^2(\ell - 2). \quad (8)$$

2. Applications

2.1. Piecewise long memory Gaussian processes

Assume that the process $X = (X_t)_{t \in \mathbb{N}}$ is a Gaussian piecewise long-range dependent (LRD) process, i.e. there exists $(D_j^*)_{0 \leq j \leq m} \in (0, 1)^{m+1}$ and for all $j = 0, \dots, m$ and $k \in \{[N\tau_j^*], [N\tau_j^*] + 1, \dots, [N\tau_{j+1}^*] - 1\}$, $X_k = X_{k - [N\tau_j^*]}^{(j)}$, where $X^{(j)} = (X_t^{(j)})_{t \in \mathbb{N}}$ satisfies the following Assumption LRD(D_j^*).

Assumption LRD(D): Y is a centered stationary Gaussian process with spectral density f such that $f(\lambda) = |\lambda|^{-D} \cdot f^*(\lambda)$ for all $\lambda \in [-\pi, \pi] \setminus \{0\}$ with $f^*(0) > 0$ and with $C_2 > 0$, $|f^*(\lambda) - f^*(0)| \leq C_2 \cdot |\lambda|^2$ for all $\lambda \in [-\pi, \pi]$.

Following [5], if the mother wavelet is supposed to be included in a Sobolev ball, then

Corollary 1 Let X be a Gaussian piecewise LRD process defined as above and $\psi : \mathbb{R} \mapsto \mathbb{R}$ be $[0, 1]$ -supported with $\psi(0) = \psi(1) = 0$ and $\int_0^1 \psi(t) dt = 0$ and such that there exists sequence $(\psi_\ell)_{\ell \in \mathbb{Z}}$ satisfying $\psi(\lambda) = \sum_{\ell \in \mathbb{Z}} \psi_\ell e^{2\pi i \ell \lambda} \in \mathbb{L}^2([0, 1])$ and $\sum_{\ell \in \mathbb{Z}} (1 + |\ell|)^{5/2} |\psi_\ell| < \infty$. Under Assumption C, for all $0 < \kappa < 2/15$, if $a_N = N^{\kappa+1/5}$ and $v_N = N^{2/5-3\kappa}$ then (5), (6), (7) and (8) hold.

Thus, the rate of convergence of $\hat{\tau}$ to τ^* (in probability) is $N^{2/5-3\kappa}$ for $0 < \kappa$ arbitrary small. Estimators \tilde{D}_j and \bar{D}_j converge to the parameters D_j^* following a central limit theorem with a rate of convergence $N^{2/5-\kappa/2}$ for $0 < \kappa$. Convincing results of simulations can be observed in Table 1 and Figure 1.

	$N = 20000$			$N = 5000$					$N = 10000$				
	$\tau_1:0.75 \ D_0:0.2 \ D_1:0.8$			$\tau_1:0.3$		$\tau_2:0.78$		$H_0:0.6$	$H_1:0.8$	$H_2:0.5$			
Estim.	$\hat{\tau}_1$	\bar{D}_0	\bar{D}_1	$\hat{\tau}_1$	$\hat{\tau}_2$	\bar{H}_0	\bar{H}_1	\bar{H}_2	$\hat{\tau}_1$	$\hat{\tau}_2$	\bar{H}_0	\bar{H}_1	\bar{H}_2
	0.7540	0.1902	0.7926	0.3465	0.7942	0.5578	0.7272	0.4395	0.3086	0.7669	0.5597	0.7633	0.4993
$\hat{\sigma}_{Estim.}$	0.0215	0.0489	0.0761	0.1212	0.1322	0.0595	0.0837	0.0643	0.0893	0.0675	0.0449	0.0813	0.0780
\sqrt{MSE}	0.0218	0.0499	0.0764	0.1298	0.1330	0.0730	0.1110	0.0883	0.0897	0.0687	0.0604	0.0892	0.0780

Table 1

Left: Estimation of τ_1 , D_0 and D_1 in the case of piecewise FARIMA(0, d_j , 0) ($d_1 = 0.1$ and $d_2 = 0.4$) with one change point when $N = 20000$ (50 realizations). Right: Estimation of τ_1 , τ_2 , H_0 , H_1 and H_2 in the case of piecewise FBM with two change points when $N = 5000$ and $N = 10000$ (50 realizations)

2.2. Piecewise fractional Brownian motions

Now, X will be called a piecewise fractional Brownian motion if there exist two families of parameters $(H_j^*)_{0 \leq j \leq m} \in (0, 1)^{m+1}$ and $(\sigma_j^{*2})_{0 \leq j \leq m} \in (0, \infty)^{m+1}$ such that for all $j = 0, \dots, m$ and $t \in [[N\tau_j^*], [N\tau_j^*] + 1, \dots, [N\tau_{j+1}^*] - 1]$, $X_t = X_{t - [N\tau_j^*]}^{(j)}$, where $X^{(j)} = (X_t^{(j)})_{t \in \mathbb{R}}$ is a FBM with parameters H_j^* and σ_j^{*2} . Following the results of [3], one obtains,

Corollary 2 Let X be a piecewise FBM and $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be a piecewise continuous and left (or right)-differentiable, such that $|\psi'(t^-)|$ is Riemann integrable with $\psi'(t^-)$ the left-derivative of ψ in t , with support included in $[0, 1]$ and $\int_{\mathbb{R}} t^p \psi(t) dt = \int_0^1 t^p \psi(t) dt = 0$ for $p = 0, 1$. Let $A := |\sup_j H_j^* - \inf_j H_j^*|$. If $A < 1/2$, under Assumption C, for all $0 < \kappa < \frac{1}{1+4A} - \frac{1}{3}$, if $a_N = N^{1/3+\kappa}$ and $v_N = N^{2/3(1-2A)-\kappa(2+4A)}$ then (5), (6), (7) and (8) hold.

Remark 3 The dependence of this result on A can be explained by the fact that $2(\sup_j \alpha_j^* - \inf_j \alpha_j^*) + 1$, with $\alpha_j^* = 2H_j^* + 1$, has to be smaller than 3 since $a_N \cdot N^{-1/3} \xrightarrow{N \rightarrow \infty} \infty$. However, Corollary 2 is quite surprising: the smaller A , i.e. the smaller the differences between the parameters H_j , the faster the convergence rates of estimators $\hat{\tau}_j$ to τ_j^* . If the difference between two successive parameters H_j is too large, the estimators $\hat{\tau}_j$ do not seem to converge. This is attributable to the influence of the other segments that is even deeper than the involved exponents are different (simulations exhibit this paroxysm in [6]).

Thus, the rate of convergence of $\hat{\tau}$ to τ^* (in probability) can be $N^{2/3(1-2A)-\kappa'}$ for $0 < \kappa'$ as small as one wants when $a_N = N^{1/3+\kappa'/(2+4A)}$. Results of simulations can be observed in Table 1 and Figure 1 in a case where $A = 0.3 < 1/2$.

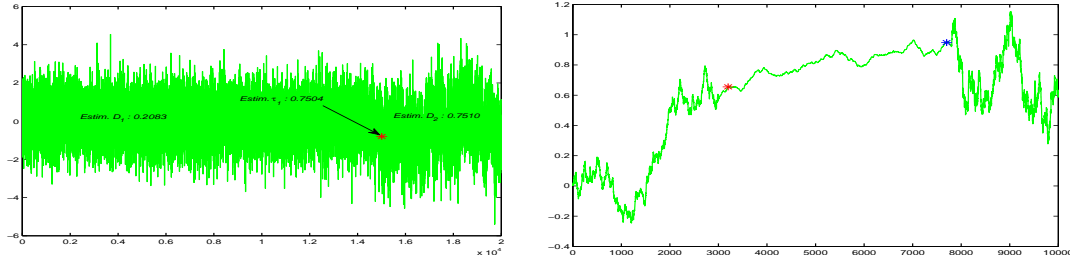


Figure 1. Left: Piecewise FARIMA(0, d_j , 0) (with $d_0 : 0.1$ ($D_0 : 0.2$), $d_1 : 0.4$ ($D_1 : 0.8$) and $\tau_1 : 0.75$). Right: Piecewise FBM(H_j) ($\tau_1 : 0.3$, $\tau_2 : 0.78$, $H_0 : 0.6$, $H_1 : 0.8$ and $H_2 : 0.5$), ($\hat{\tau}_1 : 0.32$, $\hat{\tau}_2 : 0.77$, $\hat{H}_0 : 0.5608$, $\hat{H}_1 : 0.7814$ and $\hat{H}_2 : 0.4751$).

References

- [1] Abry P., Veitch D., Flandrin P., *Long-range dependent: revisiting aggregation with wavelets* JTSA, 19, 1998, 253-266.
- [2] Ayache A., Bertrand P., Lvy Vhel J., *A central limit theorem for the quadratic variations of the step fractional Brownian motion*, Statistical Inference for Stochastic Processes 10, 2006, 1-27.
- [3] Bardet J.M., *Statistical Study of the Wavelet Analysis of Fractional Brownian Motion*, IEEE Trans. Inf. Theory, Vol. 48, No. 4, 2002, 991-999.
- [4] Bardet J.M., Bertrand P., *Identification of the multiscale fractional Brownian motion with biomechanical applications*, Journal of Time Series Analysis, 2007, 1-52.
- [5] Bardet J.M., Bibi H., Jouini A., *Adaptive wavelet based estimator of the memory parameter for stationary Gaussian processes*, to appear in Bernoulli, 2008.
- [6] Bardet J.M., Kammoun I. *Detecting changes in the fluctuations of a Gaussian process and an application to heartbeat time series*, Preprint Hal-00194909, 2007.
- [7] Beran J., Terrin N., *Testing for a change of the long-memory parameter*, Biometrika, 83, 1996, 627-638.
- [8] Flandrin P., *Wavelet analysis and synthesis of fractional Brownian motion*. IEEE Trans. Inf. Theory, 38, 1992, 910-917.
- [9] Giraitis L., Leipus R., Surgailis D., *The change-point problem for dependent observations*, Journal of Statistical Planning and Inference, 53, 1996, 297-310.
- [10] Horváth L., *Change-Point Detection in Long-Memory Processes*, Journal of Multivariate Analysis, 78, 2001, 218-134.
- [11] Kokoszka P.S., Leipus R., *Detection and estimation of changes in regime*, In P. Doukhan, G. Oppenheim and M.S. Taqqu editors, *Long-range Dependence: Theory and Applications*, Birkhäuser, 2003, 325-337.
- [12] Lavielle M., *Detection of multiple changes in a sequence of dependent variables*, Stoch. Proc. Appl., 83, 1999, 79-102.
- [13] Lavielle M., Ludea C., *The multiple change-points problem for the spectral distribution*, Bernoulli 6, 2000, 845-869.
- [14] Lavielle M., Teyssi re G. *Detecting Multiple Change-Points in Multivariate Time Series*, Lithuanian Mathematical Journal 46, 2006, 351-376.

- [15] Moulines E., Roueff F., Taqqu, M.S., *On the spectral density of the wavelet coefficients of long memory time series with application to the log-regression estimation of the memory parameter*, J. Time Ser. Anal., 2007, 155-187.